proteins are needed for rapid, efficient accumulation of the complex at the sites of DSBs. PARP1 may cooperate with the MRN complex to facilitate signaling of DSBs. CtIP (or Ctp1) is another protein that has recently been linked with MRN (Figure 2C). It is a mammalian tumor suppressor whose presence in the nucleus is limited to the S and G2 phases of the cell cycle. A recent study indicates that CtIP can form a complex with MRN, directly interacting with Nbs1 in a cell-cycle-dependent manner. The formation of this complex, which also includes BRCA1, requires cyclin-dependent kinase activity. Recent findings indicate that this Brca1–MRN–CtIP complex is important for facilitating DSB resection, which generates the 3' overhanging single-stranded DNA that is needed both for HR-mediated DSB repair and for the maintenance of checkpoint signaling.

What else is left to be examined?

Much has been discovered concerning the highly pleiotrophic functions of the MRN complex and new findings are continuously adding complexity. The detailed mechanistic understanding of how MRN really works *in vivo* remains elusive, however. The real challenge for the future is the integration of all the recent discoveries into mechanism. Therefore much remains to be done and no doubt there are several surprising discoveries still to be made.

Where can I find out more?

Chen, L., Nievera, C., Lee, A.Y., and Wu, X. (2008). Cell cycle-dependent complex formation of BRCA1/CtIP/MRN is important for DNA double-strand break repair. J. Biol. Chem. *283*, 7713–7720.

Haber, J.E. (1998). The many interfaces of Mre11. Cell *95*, 583–586.

Haince, J.F., McDonald, D., Rodrigue, A., Déry, U., Masson, J.Y., Hendzel, M.J., and Poirier, G.G. (2008). PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. J. Biol. Chem. *283*, 1197–1208.

Lavin, M.F. (2007). ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. Oncogene *26*, 7749–7758.

Lee, J.H., and Paull, T.T. (2007). Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. Oncogene *26*, 7741–7748.

Williams, R.S., Williams, J.S., and Tainer, J.A. (2007). Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template. Biochem. Cell Biol. *85*, 509–520.

Genome Stability Laboratory, Department of Biochemistry and National Centre for Biomedical Engineering Science, National University of Ireland Galway, Galway, Ireland.
*E-mail: noel.lowndes@nuigalway.ie

Primer

Evolution of human vocal production

Asif A. Ghazanfar[1] and Drew Rendall[2]

Determining the substrates required for the evolution of human speech is difficult as most traits thought to give rise to human speech — the vocal production apparatus and the brain — do not fossilize. Nor do we have any 'proto-human' sound tracks to analyse. The fossil record is also of limited utility for identifying indicators of pre-historic linguistic abilities. Ultimately, we are left with only one reliable way of investigating the biological mechanisms underlying the evolution of speech: the comparative method. By comparing the vocal behavior and biology of extant primates with humans, we can deduce the behavioral capacities of extinct common ancestors, allowing identification of homologies and providing clues as to the adaptive functions of such behaviors. Here we focus on what we have learnt about the evolution of vocal production in primates from the comparative approach.

Basic mechanisms of vocal production

In human and nonhuman primates, the anatomy and basic mechanics of voice production are broadly similar (Figure 1). Voice production involves a sound source, generally the larynx, coupled to a sound filter represented by the vocal-tract airways (the oral and nasal cavities) above the larynx. These two basic components of the vocal apparatus behave and interact in complex ways to generate a wide range of sounds. The most common and best-studied modes of vocal production involve a stable vibration of the vocal folds of the larynx. Such vibration generates a complex, but highly patterned, sound source composed of a fundamental frequency — corresponding to the base rate at which the vocal folds vibrate — and multiple harmonic overtones of the fundamental frequency (Figure 2, top panel).

This rich frequency spectrum then passes through the airways of the vocal tract above the larynx. Energy at frequencies that coincide with the natural resonance frequencies of these airways are passed easily, while energy at other frequencies is absorbed by the vocal tract walls and is thus attenuated (Figure 2, middle panel).

Ultimately, then, the sound that is radiated at the lips is, to a first approximation, a linear combination of the original, laryngeal sound source subsequently filtered by the resonance properties of the vocal tract (Figure 2, bottom panel). Among the perceptually salient dimensions of this complex sound, two in particular stand out and are directly traceable to these landmark components of vocal apparatus: voice pitch, which is the perceptual correlate of the fundamental frequency and determined by the length and mass of the vocal folds; and voice timbre, which is the perceptual correlate of the vocal tract resonances (or formants), which are determined by the length and cross-sectional area of the vocal tract airways. These two aspects of the voice feature prominently in the social communication of human and nonhuman primates.

Differences in the vocal apparatus across primates

While the essentials of vocal production are similar across primates, there are important differences between the production of human speech and of nonhuman primate vocalizations. Some of these differences can be directly attributed to anatomical changes during the course of evolution. Here we describe three: the descended larynx; increased thoracic innervation; and laryngeal air sacs.

The descended larynx

One of the most conspicuous differences in vocal anatomy between human and nonhuman primates is the descended position of the larynx in the human vocal tract relative to its position higher in the vocal tract of nonhuman primates (Figure 1). The result is, effectively, a two-tube vocal tract in humans composed of the oral cavity common to all primates, and an additional enlarged pharyngeal

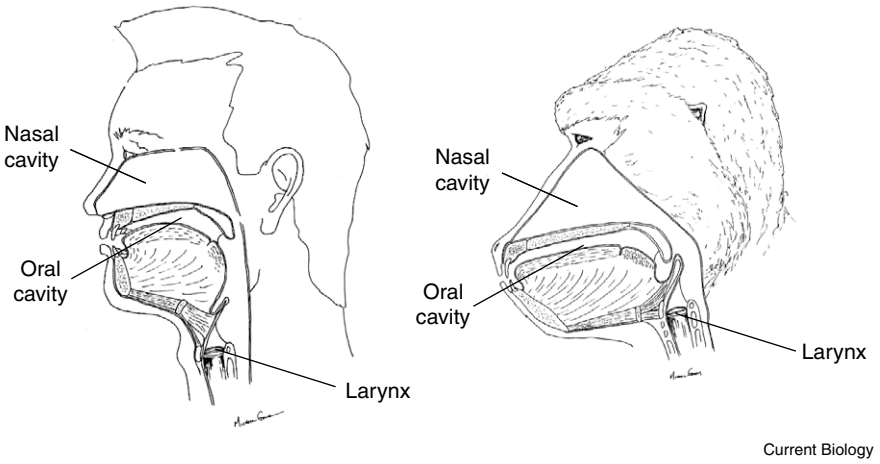

Figure 1. Anatomy of the vocal apparatus.

In both human and nonhuman primates, the source of voiced sounds involves vibration of the vocal folds of the larynx. This sound source travels up the vocal tract, where the oral and nasal cavities of the vocal tract act as a filter, passing acoustic energy at some frequencies and attenuating energy at other frequencies according to their size- and shape-specific transfer function. The sound radiated at the lips thus reflects the combined action of an acoustic source and a filter. Note that the larynx lies lower in the vocal tract of the human compared with the baboon, creating a second large cavity (the pharynx) at the back of the mouth that is relatively absent in the baboon. (Artwork by Michael Graham; adapted with permission from APA from Rendall, D. et al. (2007). J. Exp. Psychol. Human Percept. Perform. *33*, 1208–1219.)

cavity seen only in humans. This two-tube configuration, coupled with an agile tongue and a capacity for rapid mandible and lip movements, allows humans considerable articulatory latitude when vocalizing. These movements produce dynamic changes to the resonance properties of the vocal tract. Because it is exactly these dynamically changing resonances that define many of the phonemes of contemporary languages, the descent of the larynx in humans has long been considered a key anatomical adaptation for language. Indeed, its first appearance in the human fossil record is considered by some to mark the origins of language in early hominins. Unfortunately, laryngeal position is difficult to establish in fossils and so the laryngeal position of fossil species, such as Neanderthals, remains hotly debated.

At the same time, alternative functional arguments for a descended larynx have been proposed. For example, a descended larynx also produces a longer vocal tract with requisitely lower resonance frequencies and it has been argued that laryngeal descent did not evolve *for* language (even if later co-opted to it) but rather to exaggerate body size which was (and may still be) critical to social influence in human and primate societies. Another possibility is that laryngeal descent occurred simply as a by-product of the cranial remodeling that we know was producing a dramatically shorter face (and globular neurocranium) in early hominids. Without a compensatory descent of the larynx to preserve overall vocal-tract length, the foreshortened face (and vocal-tract) would have seriously disrupted all of the co-evolved systems of vocal production and perception — systems that are functionally linked to cueing the identity and physical characteristics of the speaker. While these possibilities remain speculative, their potential relevance to language origins leave them as critical unresolved issues for future research.

Increased thoracic innervation and breath control

One necessary feature for the production of fully modern human speech is the fine control of breathing. Breath control fuels sound production. Most human speech takes place on expirations alone, interspersed with rapid inspirations, in a pattern very different from the more evenly divided cycles of quiet breathing. The control of air-pressure from the respiratory system to the vocal folds also enables sound production to continue beyond the point at which the normal minimum lung volume in quiet breathing is reached. This exquisite control of breathing enables humans to speak fluently in long sentences, without disruptive pauses for inspirations, and with the necessary quick inspiratory pauses placed at meaningful linguistic boundaries.

The thoracic region of the spinal cord is involved in controlling respiration and the size of the thoracic vertebral canal gives an indication of amount of innervation in this region of the spinal cord. A comparative analysis of the thoracic vertebral canal reveals that modern humans and Neanderthals have an expanded canal when compared to other extinct hominids and extant nonhuman primates. Although there are many possible explanations for this increase in thoracic innervation, including postural control for bipedalism and respiration for endurance, comparative analyses rule out these possibilities because of their evolutionary timing or because they are insufficiently demanding neurologically. The remaining possible function is increased control of breathing for speech.

If, however, there is a direct relationship between increased thoracic innervation and the evolution of human speech, then the breathing control required for nonhuman primate vocalizations should be substantially less than that needed for human speech. Unfortunately, very little data are available on this issue as most work on primate vocal production focuses on other aspects of vocal control. The available evidence suggests that when primates produce a long sequence of sounds, it is based on a series of both expirations and inspirations. This is in contrast to human speech in which multiple units of sounds are produced during an extended expiration. In nonhuman primates, breathing rate imposes a limit on call duration and calling rate. There may, however, be another way around this limitation: air sacs.

Laryngeal air sacs

Many, but not all, primates have sac-like extensions of the larynx or other parts of the vocal tract. These air sacs are of different sizes and located in different positions

depending on the species. Humans, oddly enough, do not have air sacs (though there is some evidence of vestigial ones) and it is thought that they were lost during the course of evolution. We know very little about the function of air sacs in the primates who possess them and, thus, very little about why humans do not have them, but many hypotheses have been suggested. Most of these are related to vocal production in one form or another.

One hypothesis suggests that air sacs, at least in Old World monkeys, amplify vocalizations by acting as resonance chambers. Amplification would be important for species that produce territorial or mating calls that must travel great distances. Another hypothesis is that, like the descended larynx, the air sacs modify the calls to make them sound like they are being produced by a larger animal. A third hypothesis suggests that the entry way to the air sacs acts as a secondary kind of vocal fold and air passing through it can produce sound independently of the actual vocal folds located in the larynx. A final hypothesis proposes that the airs sacs can be used to store air that can be re-breathed to allow for very long vocal sequences.

Despite their presence in many primates, from lemurs to monkeys to great apes, there have been few experimental tests of these hypotheses. For example, the intensity of the long distance 'boom' vocalization of one forest monkey species was reduced when the air sac was experimentally punctured, lending credence to the idea that they may have a role in amplification. The notion that air sacs could allow for higher calling rates and longer bouts also has some comparative support. For example, there is a tendency for smaller-bodied primates to have faster calling rates and shorter duration calls when compared with larger primates. However, those primate species with air sacs essentially by-pass this body-size-related constraint on respiratory capacity.

The neocortical control of vocal production

One of the more puzzling differences between human and non-human primate vocal production is the apparent lack of neocortical control in the latter. In humans, electrical stimulation of motor cortex produces vowel-like sounds (phonation) and sometimes lip and facial movements (articulation), while stimulation of other parts of the human frontal neocortex can produce effects related to a mosaic of separable speech functions, and damage to these neocortical areas impairs speech. Depending on the location, lesions result in difficulties in production and fluency or outright muteness. Furthermore, the extent of impairments is dependent upon which hemisphere is damaged, with damage to the left hemisphere resulting in greater impairments.

When similar electrical stimulation approaches are applied to nonhuman primates, the results are very different. Stimulation of the motor and premotor cortex (presumptive homologues of Broca's region) in both great apes and monkeys results in face, tongue and vocal cord movements, but no overt vocal production. When experimental lesions are made in these areas, the outcome is not very dramatic. In monkeys, lesions of these motor cortical areas occasionally produce a weakened voice in spontaneous vocalizations and have little or no affect on trained vocal behaviors. These findings have led investigators to suggest that, in stark contrast to humans, the frontal cortex of nonhuman primates has little or no role in the production of vocalizations and, by extension, little voluntary control of their communicative behavior.

Unfortunately, this topic has not been pursued in earnest for decades now, but there is a strong sense among many investigators that it needs to be revisited. This is in part because we have greater insights into the vocal behavior of monkeys and apes, which suggest that their vocalizations are not mindless emotional responses, and a greater knowledge of motor cortical circuitry and function. We also have a better understanding of primate vocal production mechanisms (see above), particularly with regard to vocal tract resonances, which suggest that the effects of electrical stimulation and lesions may be more subtle than simply eliciting vocalizations in the former or eliminating vocalizations in the latter. Indeed, it is very likely that these experimental manipulations have a large effect on the articulatory motions — lip, jaw and possibly tongue movements — related to producing species-typical

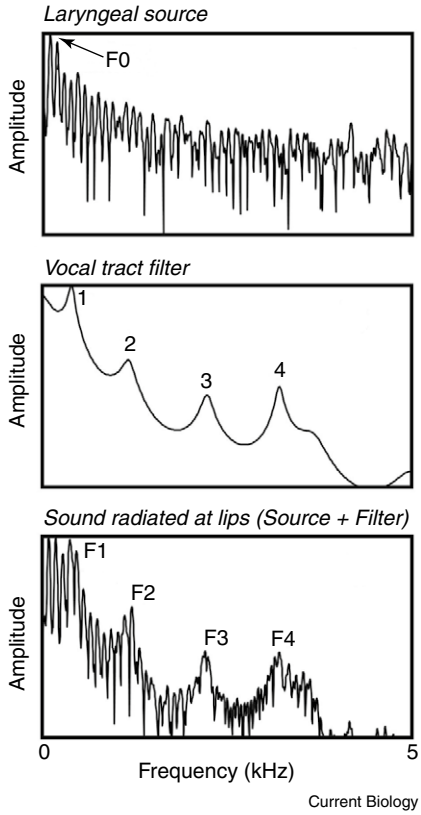

Figure 2. Voice acoustics.

The frequency spectrum of voiced sounds at the laryngeal source is shown in the top panel, where the periodicity of vocal-fold vibration is manifest by the very regular spacing of peaks. The fundamental frequency of vocal fold vibration (*F0)* appears as the lowest peak visible in this spectrum (marked by the arrow) and as the constant distance between all successive peaks, which are simply integer multiples, or harmonics, of the *F0*. The transfer function of the vocal-tract cavities is shown in the middle panel, where the numbered peaks (1–4) correspond to the frequencies of the first four resonances of the vocal tract (which in this case are for the neutral, or schwa, vowel spoken by a man). The bottom panel shows the frequency spectrum of the sound that is radiated at the lips which is approximately a linear combination of the laryngeal source spectrum passed through the vocal-tract transfer function. Therefore, it retains the many regularly-spaced harmonic peaks of the *F0*, but it now shows a more complex global frequency envelope shaped by the vocal-tract filter, where the four broad peaks (*F1–F4*) reflect the first four resonances of this filter and are typically referred to in human speech as *formants*. (From Rendell *et al.* J. Exp. Psychol. *33*, 1208–1209; adapted with permission from American Psychological Association.)

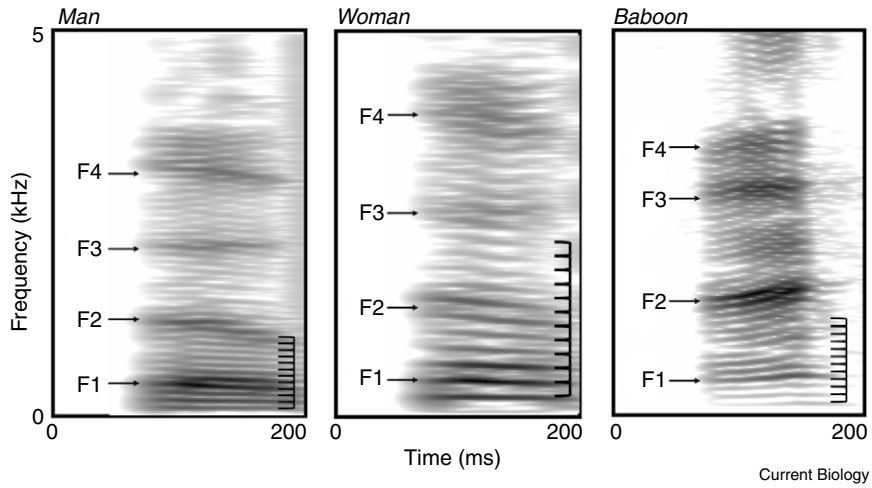

Figure 3. Spectrograms of voiced sounds.

The first two panels show spectrograms (frequency-by-time graphics) of a single vowel sound spoken by an adult male and an adult female speaker. The third panel shows a spectrogram of a vowel-like grunt vocalization produced by a baboon. In each panel, voice *F0* is manifest as the lowest single frequency band in the spectrogram as well as by the closely spaced harmonics of it (indicated by the vertically oriented grid). The resonances of the vocal tract (*F1–F4*) are manifest as broad dark bands (smudges) each of which incorporates one or more harmonics. The first two panels illustrate basic sex differences in voice pitch and formants in humans, while the last panel illustrates the obvious continuity in voiced sounds produced by humans and some nonhuman primates stemming from similarity in the basic anatomy and mechanics of voice production in the two groups.

vocalizations. These factors have a strong influence on the acoustic features of calls — features that influence the behavior of conspecifics.

The link between production and perception

In both human and nonhuman primates, the fundamental frequency and vocal tract resonances of the voice can carry a rich set of indexical cues that effectively 'point out' the caller in various ways: by their sex, relative age or body size, and individual identity. These cues arise inevitably from developmental and individual differences in the shape of the vocal folds and vocal tract. These indexical cues are perceptually salient to listeners and can be extremely useful in many social contexts. For example, among social primates, the dynamics of daily social interactions are heavily influenced by the sex, age, kinship and individual identity of interacting agents, as well as their relative positions in the social hierarchy. Studies of numerous species have shown that the vocalizations that are used in these social interactions often contain cues to caller identity and that listeners are sensitive to these cues. Many of these same social dimensions underpin routine human social interactions and are cued by the same voice features.

Recently, there has been growing interest in other constitutional dimensions of signalers that might be conveyed through the same voice cues. For example, in both human and nonhuman primates, body size can be an important determinant of social interactions, with larger, socially-mature individuals enjoying greater privilege generally. In some cases, social interactions are mediated vocally. Here, it is intuitive that larger individuals would be marked by lower fundamental and resonance frequencies because they would, all else being equal, be expected to have both a larger larynx (with longer vocal folds that naturally vibrate at lower frequencies) and a longer vocal tract (with lower resonances), just as they generally have larger hands and feet and longer limbs (Figure 3). Research has confirmed some of these intuitions. For example, in some nonhuman primates, immature animals have both higher voice pitch and higher resonances than their larger-bodied adult counterparts; adult males are often larger than females and have lower F0 and resonances as well.

The same pattern of age-sex differences in voice applies to humans, but some of the voice differences exceed what would be expected from body size differences alone. For example, human males are, on average, only 10% taller and 20% heavier than females; yet, there is a twofold difference in pitch between the sexes leading to the proposal that the hypertrophied larynx of human males and their disproportionately low-pitched voices are an adaptation to exaggerate body size. There has also been more effort to test the relationships within age-sex classes in humans where there proves to be no reliable relationship between voice pitch and adult body size in either men or women. At the same time, perceptual studies show a consistent tendency for listeners to label deeper-voiced speakers as being larger. Somewhat paradoxically, however, listeners base their size assessments on the unreliable voice pitch cues rather than the more reliable vocal tract resonances and so are often very wrong in their size estimates. This paradox remains unresolved.

Further reading

Fitch, W.T., and Hauser, M.D. (1995). Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on "honest" advertisement. Am. J. Primatol. *37*, 179–190.

Hewitt, G., MacLarnon, A., and Jones, K.E. (2000). The function of laryngeal air sacs in primates: a new hypothesis. Folia Primatologica *73*, 70–94.

Ghazanfar, A.A., Turesson, H.K., Maier, J.X., van Dinther, R., Patterson, R.D., and Logothetis, N.K. (2007). Vocal tract resonances as indexical cues in rhesus monkeys. Curr. Biol. *17*, 425–430.

Lieberman, P. (1968). Primate vocalizations and human linguistic ability. J. Acoust. Soc. Am. *44*, 1574–1584.

MacLarnon, A.M., and Hewitt, G.P. (1999). The evolution of human speech: the role of enhanced breathing control. Am. J. Phys. Anthropol. *109*, 341–363.

Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. J. Acoust. Soc. Am. *117*, 944–955.

[1]Neuroscience Institute and Department of Psychology, Princeton University, Princeton, New Jersey 08540, USA. [2]Behaviour and Evolution Research Group, Department of Psychology, University of Lethbridge, Lethbridge, Alberta, Canada TIK 3M4.
E-mail: asifg@princeton.edu; d.rendall@uleth.ca